

Armed Services Technical Information Agency

Because of our limited supply, you are requested to return this copy WHEN IT HAS SERVED YOUR PURPOSE so that it may be made available to other requesters. Your cooperation will be appreciated.

AD

29564

NOTICE: WHEN GOVERNMENT OR OTHER DRAWINGS, SPECIFICATIONS OR OTHER DATA ARE USED FOR ANY PURPOSE OTHER THAN IN CONNECTION WITH A DEFINITELY RELATED GOVERNMENT PROCUREMENT OPERATION, THE U. S. GOVERNMENT THEREBY INCURS NO RESPONSIBILITY, NOR ANY OBLIGATION WHATSOEVER; AND THE FACT THAT THE GOVERNMENT MAY HAVE FORMULATED, FURNISHED, OR IN ANY WAY SUPPLIED THE SAID DRAWINGS, SPECIFICATIONS, OR OTHER DATA IS NOT TO BE REGARDED BY IMPLICATION OR OTHERWISE AS IN ANY MANNER LICENSING THE HOLDER OR ANY OTHER PERSON OR CORPORATION, OR CONVEYING ANY RIGHTS OR PERMISSION TO MANUFACTURE, USE OR SELL ANY PATENTED INVENTION THAT MAY IN ANY WAY BE RELATED THERETO.

Reproduced by
DOCUMENT SERVICE CENTER
KNOTT BUILDING, DAYTON, 2, OHIO

UNCLASSIFIED

Reproduced
FROM LOW CONTRAST COPY.

AD No. 29 564
ASTIA FILE COPY

SELECTION OF A DISTRIBUTION FUNCTION TO MINIMIZE
AN EXPECTATION SUBJECT TO SIDE CONDITONS

BY

HERMAN CHERNOFF AND STANLEY REITER

TECHNICAL REPORT NO. 23

MARCH 12, 1954

This work was sponsored by the Army, Navy, and Air
Force through the Joint Services Advisory Committee
for Research Groups in Applied Mathematics and
Statistics by Contract No. N6onr 25140 (NR-342-022)

APPLIED MATHEMATICS AND STATISTICS LABORATORY
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

Selection of a Distribution Function to Minimize
an Expectation Subject to Side Conditions

By

Herman Chernoff and Stanley Reiter
Stanford University

I. The following problem arises in a bioassay problem. We wish to choose a c.d.f. F of x so as to minimize E where,

$$E = \int g(x) dG(x)$$

subject to

$$(1) \quad \int x dG(x) = c_1$$

$$(2) \quad \int x^2 dG(x) = c_2$$

where c_1 and c_2 are given positive constants $c_2 \geq c_1^2$ and

$$g(x) = 1 - e^{-\beta x}.$$

We are also interested in finding that c.d.f. which makes E a maximum, subject to the constraints above.

We obtain the following results.

In Section II it is shown that the minimizing c.d.f. exists and that for this c.d.f. the probability is concentrated at two points, one of which is zero. The minimizing c.d.f. is,

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 - \frac{c_1^2}{c_2} & \text{for } 0 \leq x < \frac{c_2}{c_1} \\ 1 & \text{for } x \geq \frac{c_2}{c_1} \end{cases} .$$

In Section III we find that the maximizing distribution does not exist, but that the maximum can be approximated as closely as we like by a two point distribution satisfying the constraints. The approximation of the maximum to within $\epsilon > 0$ is obtained by placing very large probability at $c_1 - \delta(\epsilon)$ and the remaining probability at a value of x sufficiently large to satisfy the conditions on the expected value of x^2 .

In Section IV we compute the value of the minimum for various values of the parameters.

We find those values of $-\beta x$ for which

$$1 - e^{-\beta x} \text{ is equal to } .90, .85, .80, \dots, .01$$

respectively. For each value of $-\beta x$ so obtained we compute $\int g dF$ as a function of the coefficient of variation. The functions are plotted on the accompanying graphs.

II. A. We consider the following problem: given g_1, \dots, g_n continuous real valued functions defined on the interval $[a, b]$ we wish to find that c.d.f. F , which minimizes

$$(1.1) \quad \int g_1 dG$$

subject to the constraints

$$(1.2) \quad k_i \leq \int g_i dG \leq K_i \quad i = 2, \dots, n$$

for $G \in \mathcal{F}_{[a, b]}$, the class of c.d.f.'s on $[a, b]$.

We shall show that the minimizing c.d.f. is discrete, concentrating probability on at most n points of the interval $[a, b]$.

Proposition 1. The class $\mathcal{F}_{[a, b]}$ is convex and compact in the topology of convergence in distribution. [The compactness of Proposition 1 is a restatement of Helly's theorem].

Let the mapping $T: \mathcal{F}_{[a, b]} \rightarrow R_n$ be given by $T \cdot G = (\int g_1(x) dG(x), \int g_2 dG, \dots, \int g_n dG)$ for $G \in \mathcal{F}_{[a, b]}$, when R_n is n -dimensional Euclidean space.

Denote by Y the image of $\mathcal{F}_{[a, b]}$ under T . We shall sometimes denote points of Y by y . The transformation T is linear and therefore is continuous and preserves convexity. Hence we have:

Proposition 2. The set Y is convex, closed and bounded.

The restrictions (1.2) define a closed convex subset of Y ; call this set Y^1 .

We restate our problem as follows. We wish to find that point y in Y^1 whose first coordinate is a minimum. Since Y^1 is closed and bounded a minimizing point exists and is a boundary point of Y^1 and also of Y so long as Y^1 is non-null.

We characterize the minimum as follows. It is assumed at a boundary point \tilde{y} of Y . Hence there is a supporting hyperplane H for Y containing \tilde{y} . But H is at most $n-1$ -dimensional. Therefore y can be written as a convex combination of at most n extreme points of Y which lie in H . But it is easy to see that all extreme points of Y correspond to one point distributions, i.e. points of the form,

$$(\int g_1(x) d\mathcal{S}_c(x), \dots, \int g_n(x) d\mathcal{S}_c(x)) = (g_1(c), \dots, g_n(c))$$

where

$$\mathcal{S}_c(x) = \begin{cases} 0 & \text{for } x < c \\ 1 & \text{for } x \geq c \end{cases}$$

Thus y may be given by

$$\begin{aligned} \tilde{y} &= \sum_{j=1}^n \lambda_j (\int g_1(x) d\mathcal{S}_{c_j}(x), \dots, \int g_n(x) d\mathcal{S}_{c_j}(x)) \\ &= \sum_{j=1}^n \lambda_j (g_1(c_j), \dots, g_n(c_j)) \end{aligned}$$

where

$$\sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0.$$

From which it follows that

$$F = \sum_{j=1}^n \lambda_j \mathcal{S}_{c_j}(x)$$

is the minimizing distribution.

B. Consider the following problem. Let

$$\begin{aligned}g_1(x) &= 1 - e^{-\beta x} \\g_2(x) &= x \\g_3(x) &= x^2\end{aligned}$$

be defined for $x \geq 0$.

Let \mathcal{F} be the class of c.d.f.'s on $[0, \infty)$. We wish to find that c.d.f. F in \mathcal{F} which minimizes

$$\int g_1 dF$$

subject to

$$\int g_2 dF = c_1$$

and

$$\int g_3 dF = c_2$$

where $c_2 \geq c_1^2$.

In order to apply the results of A we shall first consider a modification of this problem as follows.

Let g_1, g_2, g_3 be defined for $x \in [0, b]$ for some $b > 0$, and let \mathcal{F} be replaced by $\mathcal{F}_{[0, b]}$, the class of c.d.f.'s on $[0, b]$. In this modification the results of II assure us that the minimum exists and that it is given by a discrete distribution concentrated at not more than three points.

In the present case this result can be sharpened. The minimizing c.d.f. is concentrated at just two points, one of which is zero. The following argument establishes this.

We know that there is a supporting plane of Y at the minimum y , that is, for some $\lambda_1, \lambda_2, \lambda_3$ and k .

$$\lambda_1 y_1 + \lambda_2 y_2 + \lambda_3 y_3 = k$$

$$\lambda_1 y_1 + \lambda_2 y_2 + \lambda_3 y_3 \geq k \quad \text{for all } y \in Y.$$

Since the first coordinate y_1 is being minimized we know that $\lambda_1 \neq 0$, and may normalize so as to make $\lambda_1 = 1$. Also, y can be written as a convex combination of extreme points of Y satisfying the (same) linear relation. We ask, "For how many x 's in $[0, b]$ is the function

$$(1) \quad g(x) + \lambda_2 x + \lambda_3 x^2$$

minimized?" We show that there are at most two, one of which is zero.

Differentiating with respect to x we find that the derivative,

$$(2) \quad \beta e^{-\beta x} + \lambda_2 + 2\lambda_3 x = 0$$

is a convex function and has at most two roots. These cannot both give minima, for if they did, there must be a maximum of (1) between them and hence a third root of (2). Thus, the case of two roots in the interior of $[0, b]$ corresponding to minima is excluded.

The following cases remain.

Consider the possibility that the function (1) is minimized at three points $0, x, b$, $0 < x < b$.

We exclude this by noting that there would then be a relative maximum of (1) between 0 and x and hence a root of (2) between 0 and x , and similarly, between x and b . This implies that there are three distinct roots of (2), which is false.

We are left with three cases; First that the minimum is assumed at just one point, which is the case when $c_2 = c_1^2$. In this case the minimizing distribution is $\delta_{c_1}(x)$ and the minimum is $(1 - e^{-\beta c_1})$.

There remain two possibilities, that the minimum involves the pair of points 0, and x or else that it involves the pair x and b . We shall show that there is a value of b , say b_0 , so that for $b > b_0$ the minimum does not involve b .

Suppose $b > c_1$. (If $b < c_1$ the restrictions cannot be satisfied.) Suppose we use two points, 0, x , with probabilities $1 - p$ and p respectively. The restrictions are then sufficient to determine x and p . We obtain

$$\begin{aligned} p x &= c_1 \\ p x^2 &= c_2 \\ x &= \frac{c_2}{c_1} \\ p &= \frac{c_1^2}{c_2} \end{aligned}$$

So that the value of $E = \int g_1 dF$ obtained when 0 is involved is

$$(1 - e^{-\beta \frac{c_2}{c_1}}) \frac{c_1^2}{c_2}$$

From strict concavity of g_1 it follows that

$$(1 - e^{-\beta \frac{c_2}{c_1}}) \frac{c_1^2}{c_2} < (1 - e^{-\beta c_1}) .$$

Let

$$\eta = (1 - e^{-\beta c_1}) - (1 - e^{-\beta \frac{c_2}{c_1}}) \frac{c_1^2}{c_2} .$$

Given $\varepsilon > 0$ we may select $b_0(\varepsilon)$ such that

$$b_0 > 1 \quad \frac{c_2}{b_0} < \varepsilon .$$

Then if we use x and b with probabilities $1 - p$ and p respectively we find that where $b > b_0$

$$(1 - p) x^2 + p b^2 = c_2$$

implies

$$p b < \frac{c_2}{b} < \frac{c_2}{b_0} < \varepsilon$$

$$p < \frac{c_2}{b^2} < \varepsilon$$

$$1 - p > 1 - \varepsilon .$$

From the first restriction it follows that x is forced to be close to c_1 , i.e.

$$(1 - p) x + p b = c_1$$

or,

$$x - c_1 = c_1 \left(\frac{1}{1-p} - 1 \right) - \frac{p b}{1-p}$$

$$|x - c_1| \leq c_1 \left(\frac{\varepsilon}{1-\varepsilon} \right) + \frac{\varepsilon}{1-\varepsilon}$$

$$\leq (c_1 + 1) \frac{\varepsilon}{1-\varepsilon} .$$

Thus x can be made to differ arbitrarily little from c_1 as $\varepsilon \rightarrow 0$.

Now since $g_1(x)$ is monotone increasing

$$p g_1(x) + (1-p) g_1(b) \geq g_1(x) > g_1(c_1) - \eta$$

for ε sufficiently small, say $\varepsilon < \varepsilon_0$. Thus, we find that using 0 and x gives a lower value of E than that obtained using x and b for $b > b_0(\varepsilon_0)$.

Thus, the minimum is attained by a c.d.f. concentrating probability $(1 - \frac{c_1^2}{c_2^2})$ at 0 and probability $\frac{c_1^2}{c_2^2}$ at $x = \frac{c_2}{c_1}$. The minimum is then $(1 - e^{-\beta \frac{c_2}{c_1}}) \frac{c_1^2}{c_2^2}$ which is independent of b .

C. We have shown in B that for the problem with $[0, \infty)$ replaced by $[0, b]$, the minimizing c.d.f. exists, that it is a distribution concentrated at 0 and one other point and that the minimum is

$(1 - e^{-\beta \frac{c_2}{c_1}}) \frac{c_1^2}{c_2^2}$, which is independent of the bound b .

We now show that this is in fact the minimum over all G&J.

Take any $G \in \mathcal{F}$ satisfying

$$\int x \, dG = c_1$$

$$\int x^2 \, dG = c_2 .$$

Then $\int g_1 \, dG$ exists.

Given $\varepsilon > 0$ we can find b so large that

$$\int_b^\infty x \, dG < \varepsilon$$

$$\int_b^\infty x^2 \, dG < \varepsilon$$

$$\int_b^\infty g_1 \, dG < \varepsilon .$$

Thus,

$$c_1 - \varepsilon < \int_0^b x \, dG = c_1^* < c_1$$

$$c_2 - \varepsilon < \int_0^b x^2 \, dG = c_2^* < c_2$$

and by the previous result,

$$(1 - e^{-\beta \left(\frac{c_2^*}{c_1^*}\right)}) \frac{c_1^*}{c_2^*} \leq \int_0^b g_1 \, dG < \int_0^\infty g_1 \, dG .$$

Since this is true for every $\epsilon > 0$ it follows by continuity that

$$(1 - e^{-\beta \frac{c_2}{c_1}}) \frac{c_1^2}{c_2} \leq \int_0^\infty g \, dG \quad \text{for all } G \in \mathcal{F}$$

and hence that the two point distribution yields the minimum.

III. We consider the problem to find that c.d.f. $F(x)$ which maximizes E .

$$E = \int (1 - e^{-\beta x}) \, dF(x)$$

subject to the constraints

$$(1) \quad \int x \, dF(x) = c_1$$

$$(2) \quad \int x^2 \, dF(x) = c_2$$

where c_1 and c_2 are given positive constants, $c_2 > c_1^2$

We notice first that if we ignore the constraint (2) then the maximizing c.d.f. is

$$\delta_{c_1}(x) = \begin{cases} 0 & \text{for } x < c_1 \\ 1 & \text{for } x \geq c_1 \end{cases}$$

This follows readily from the fact that $1 - e^{-\beta x}$ is a concave function.

Imposing an additional constraint, namely (2) cannot increase the

maximum. However, we notice that we can get as close as we please to the maximum, i.e., given $\delta > 0$ we can find x and p such that

$$(1-p)(c_1 - \delta) + p x = c_1$$

$$(1-p)(c_1 - \delta)^2 + p x^2 = c_2$$

and given $\varepsilon > 0$ we may take $\delta > 0$ such that

$$(1 - e^{-\beta(c_1 - \delta)})(1-p) + (1 - e^{-\beta x})p > (1 - e^{-\beta c_1}) - \varepsilon.$$

Thus, the second constraint requires at least a two point distribution, while the concavity of g_1 precludes actually attaining $g_1(c_1)$ with such a distribution. Hence the maximizing distribution does not exist. The maximum may be approximated as closely as we like by a two point distribution satisfying both constraints.

IV. Computational Results.

Let

$$c_1 = \bar{x}$$

$$c_2 = \bar{x}^2 + s^2$$

$$\beta \bar{x} = \gamma, \quad v^2 = \frac{s^2}{\bar{x}^2}$$

then

$$(1 - e^{-\beta \frac{c_2}{c_1}}) \frac{c_1}{c_2} = \frac{1 - e^{-\gamma(1+v^2)}}{1 + v^2} .$$

We shall consider the values of \bar{x} and β for which E is .90, .85, .80, ..., .05, .01, i.e., those values of γ for which

$$F(\gamma) = 1 - e^{-\gamma} = .90, .85, .80, \dots, .05, .01 .$$

For each value of γ so obtained we plot in the accompanying graphs the expected value E

$$E = \frac{1 - e^{-\gamma(1+v^2)}}{1 + v^2}$$

as a function of the coefficient of variation, v^2 , for values $0 \leq v^2 \leq 6$.

It is interesting that when the expectation corresponding to a value of γ is small, the effect of increasing the coefficient of variation is also small. For example, if γ is such that E is .15 then E is reduced to .098 if the coefficient of variation is increased from 0 to 6.

In the accompanying graphs we also compare E as a function of the coefficient of variation of x when the minimizing distribution is used with the corresponding quantities when x has the Gamma distribution.

Suppose that x has a Gamma distribution, i.e., the density f of x is given by

$$f(x) = \frac{\alpha^n x^{n-1} e^{-\alpha x}}{\Gamma(n)} \quad x > 0 .$$

Then

$$E(x) = \frac{n}{\alpha}$$

$$E(x^2) = \frac{n(n+1)}{\alpha^2} .$$

The coefficient of variation is given by

$$v^2 = \frac{\frac{n(n+1)}{\alpha^2} - \frac{n^2}{\alpha^2}}{\frac{n^2}{\alpha^2}} = \frac{1}{n} .$$

We also have

$$n = \frac{1}{v^2}$$

$$\alpha = \frac{1}{v^2 E(x)} .$$

Suppose we are given a mean $E(x)$ for which γ yields expectation \underline{F} .

Then

$$F = 1 - e^{-E(x)\beta} .$$

Now we assign a coefficient of variation v^2 . Then the value of F is

$$E[1 - e^{-\beta x}] = 1 - \int_0^{\infty} \frac{e^{-\beta x} \alpha^n x^{n-1} e^{-\alpha x}}{\Gamma(n)} dx$$

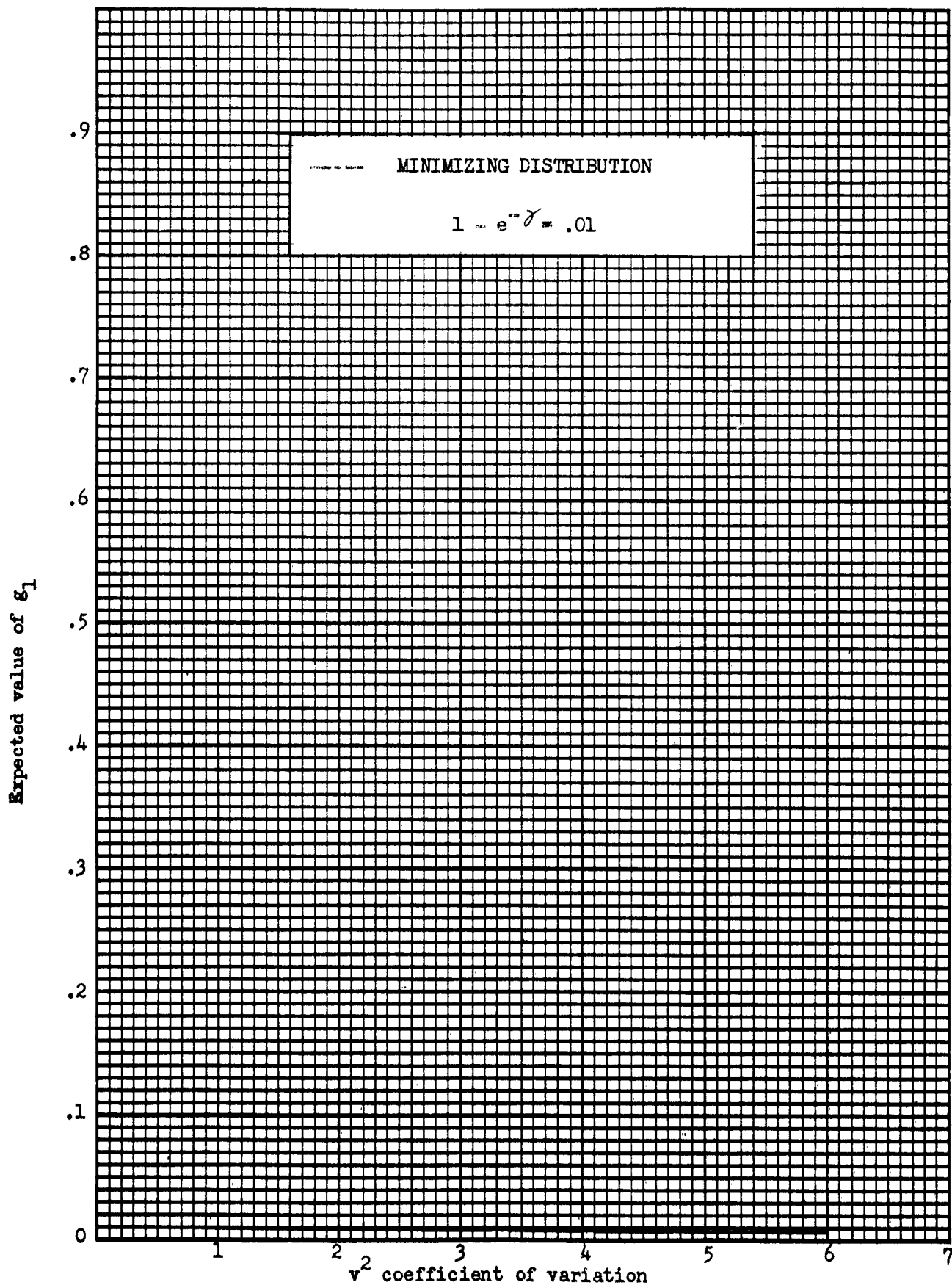
$$= [1 - (\frac{\alpha}{\alpha + \beta})^n] = 1 - [1 + \frac{\beta}{\alpha}]^{-n}$$

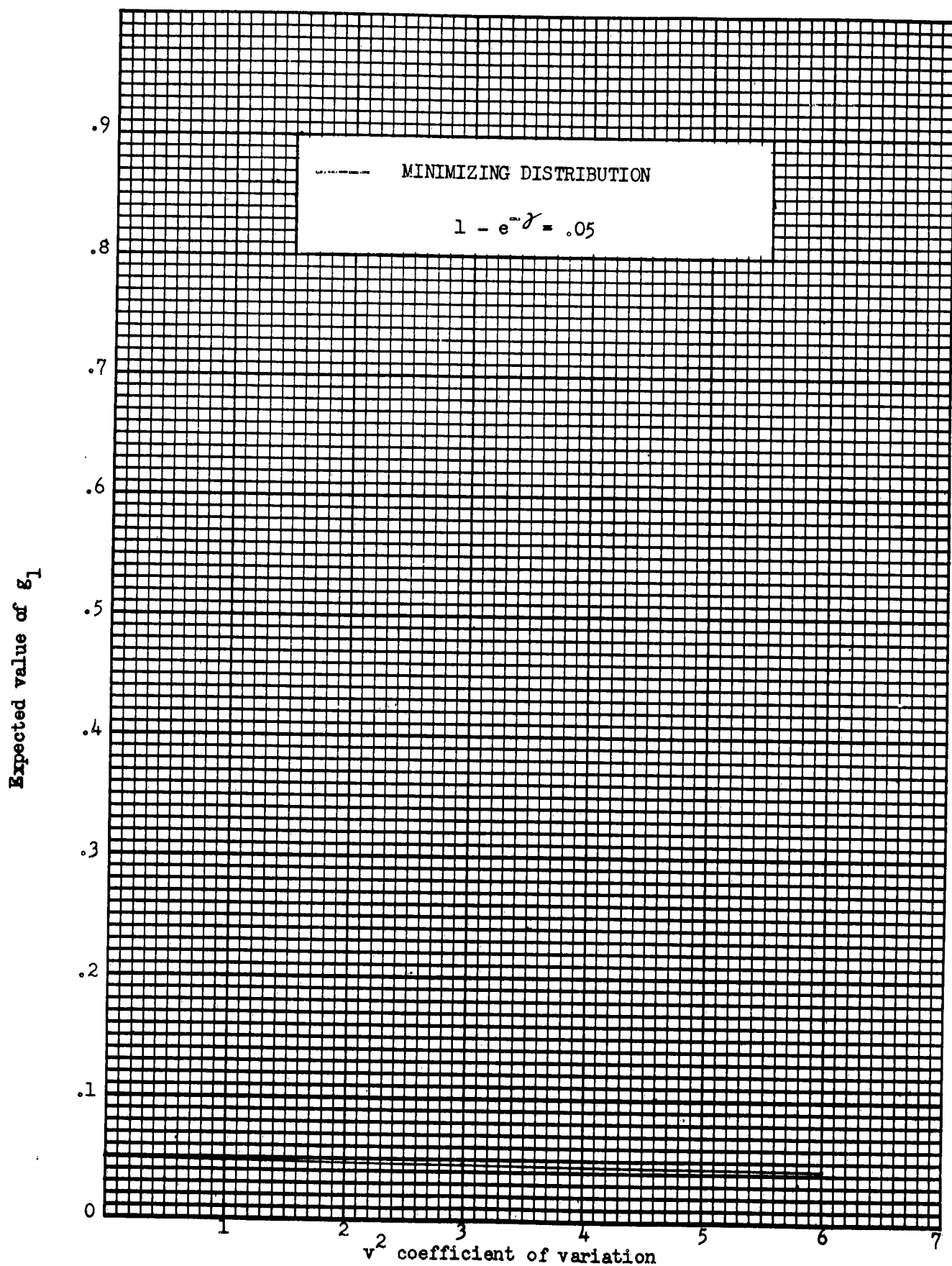
$$= 1 - \{1 + \beta E(x) v^2\}^{-\frac{1}{v^2}}$$

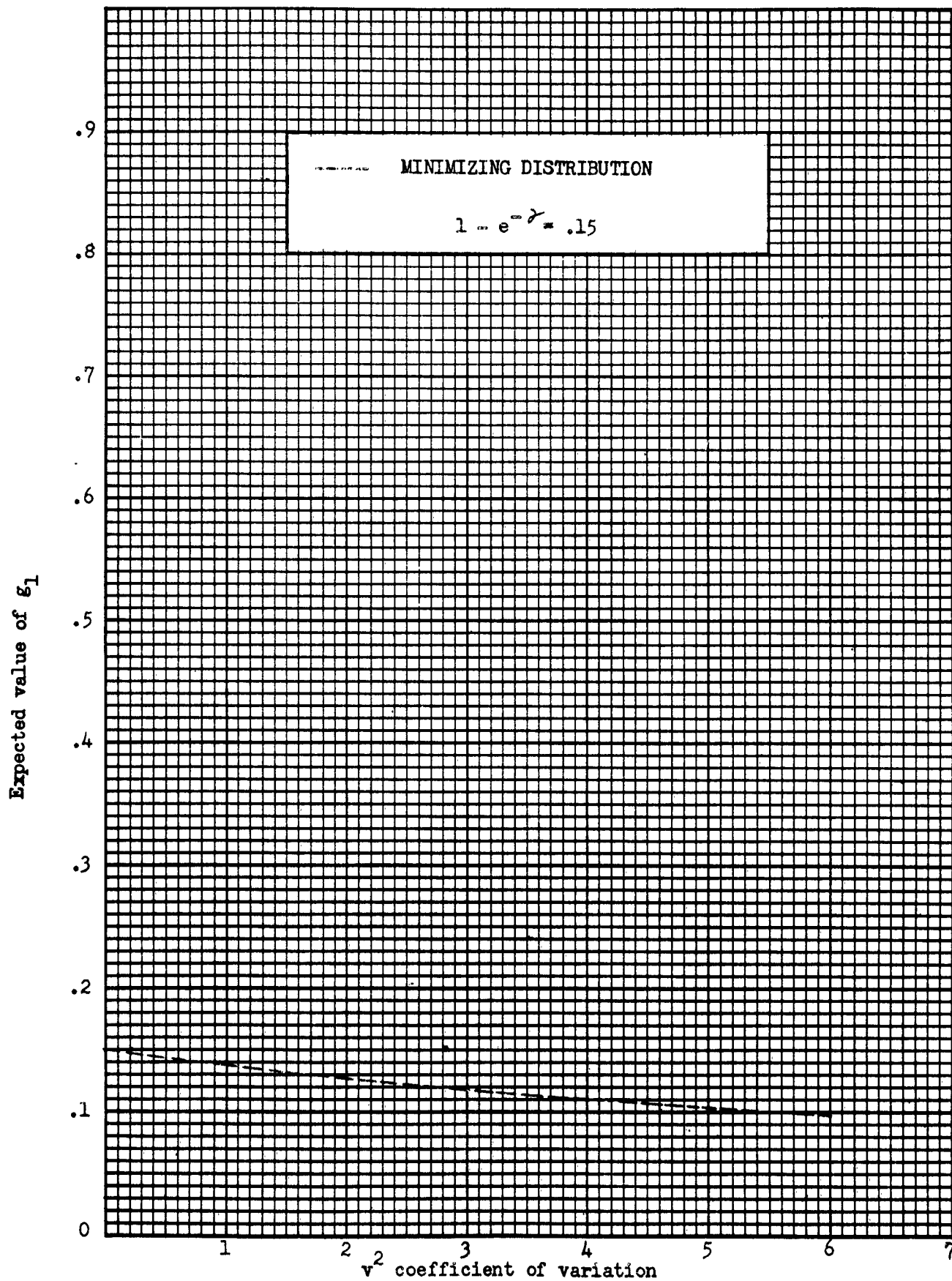
where $\beta E(x) = -\log(1 - F)$

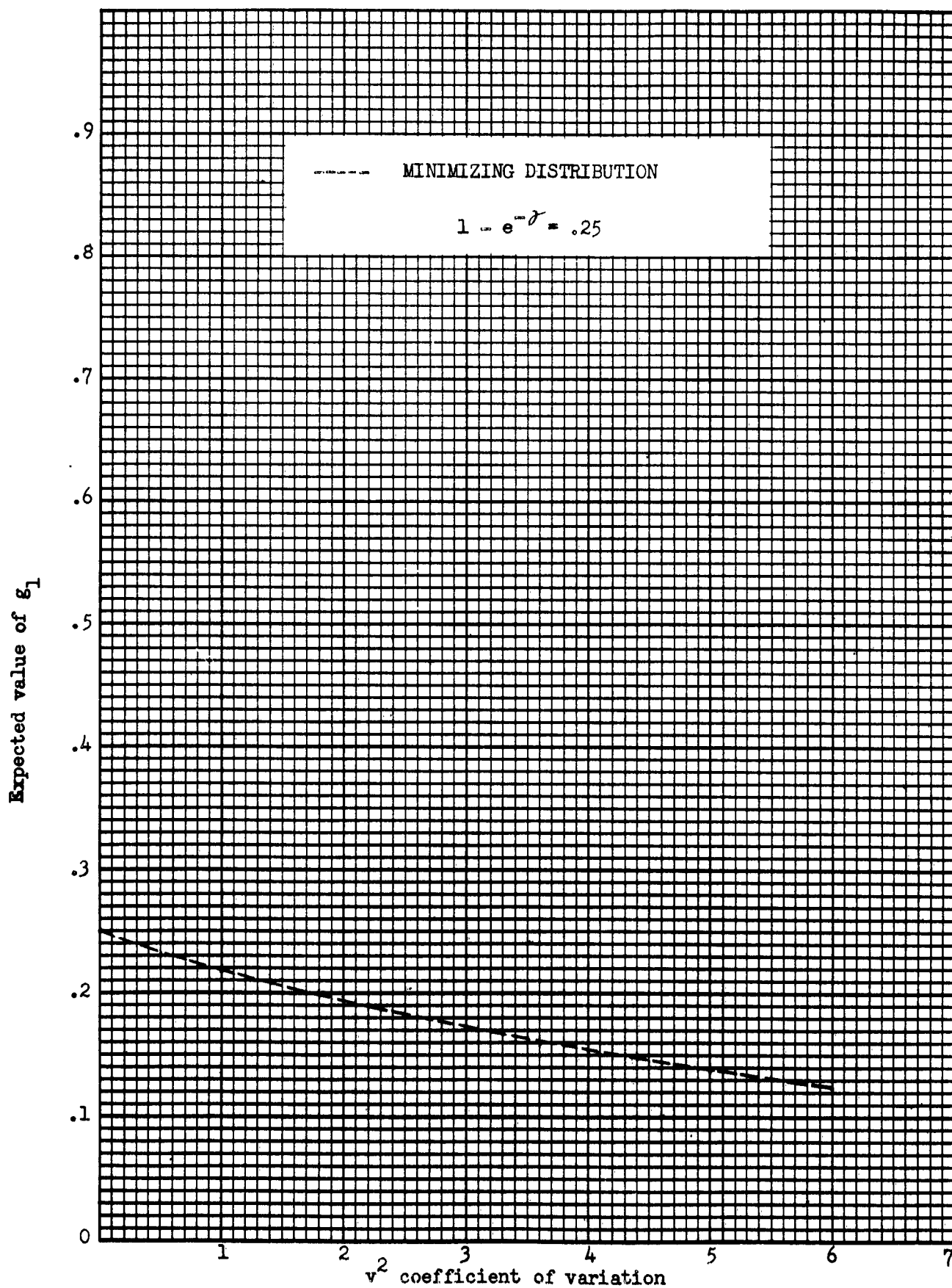
$$E[1 - e^{-\beta x}] = 1 - \{1 - v^2 \log(1 - F)\}^{-\frac{1}{v^2}} .$$

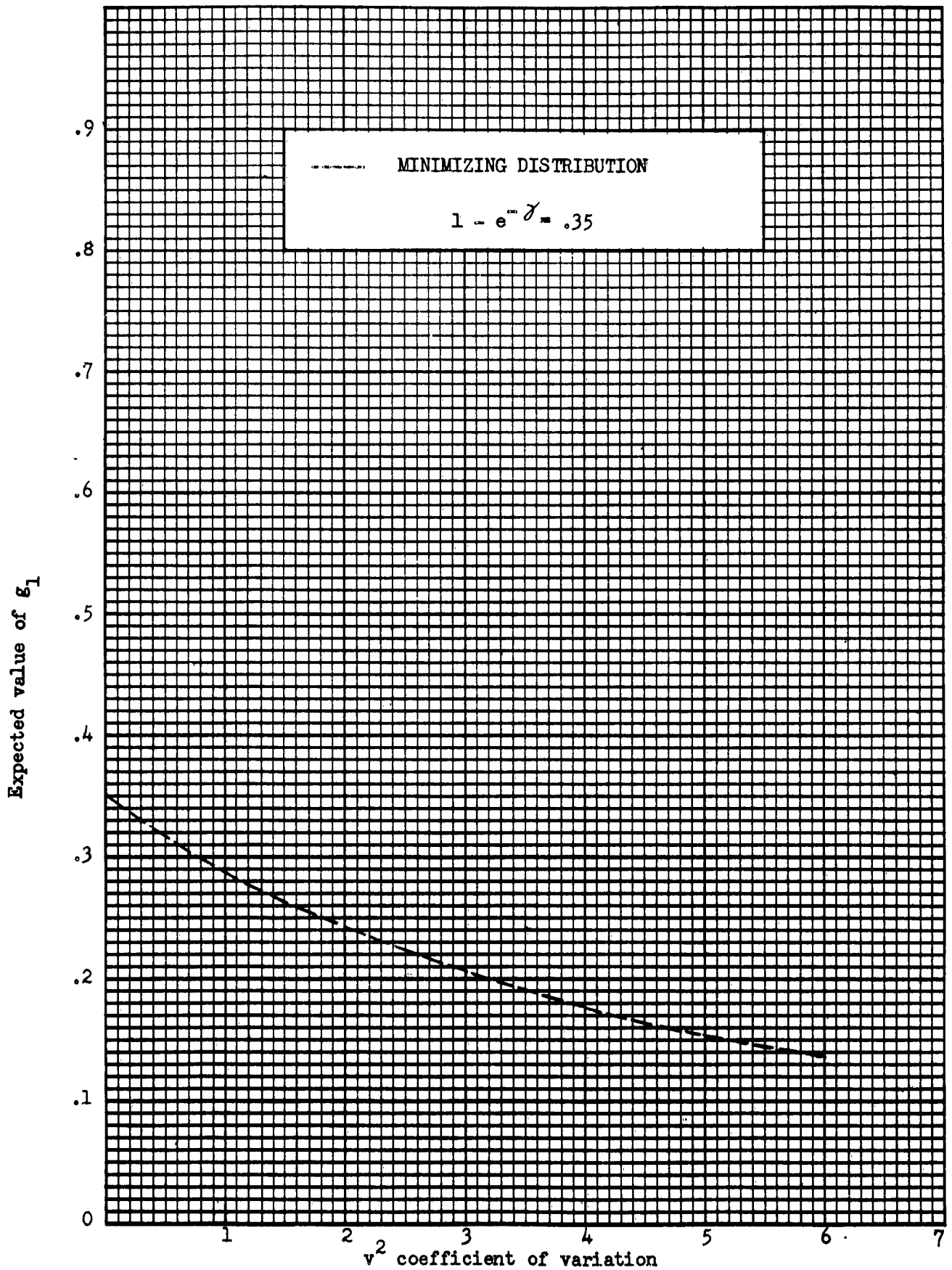
V. The methods and results of this paper are similar to those of P. C. Rosenbloom, "Quelques classes de problèmes extrémaux," Bull. Soc. Math., France 79 (1951) 1-58, 80 (1952) 183-215 and S. Karlin and L. S. Shapley, "Geometry of Moment Spaces," Memoirs of the Amer. Math. Soc., No. 12, 1953, pp. 1-93.

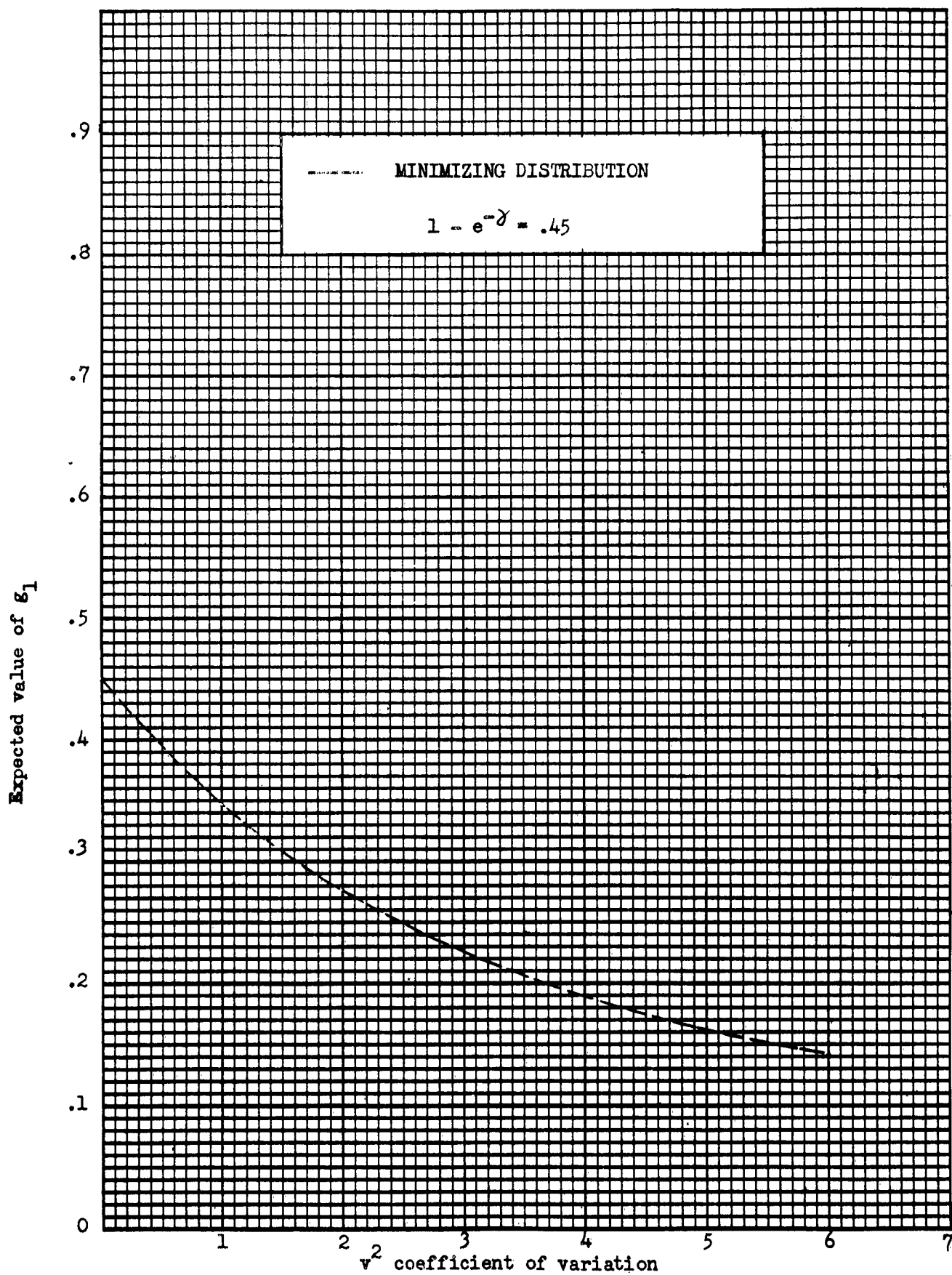


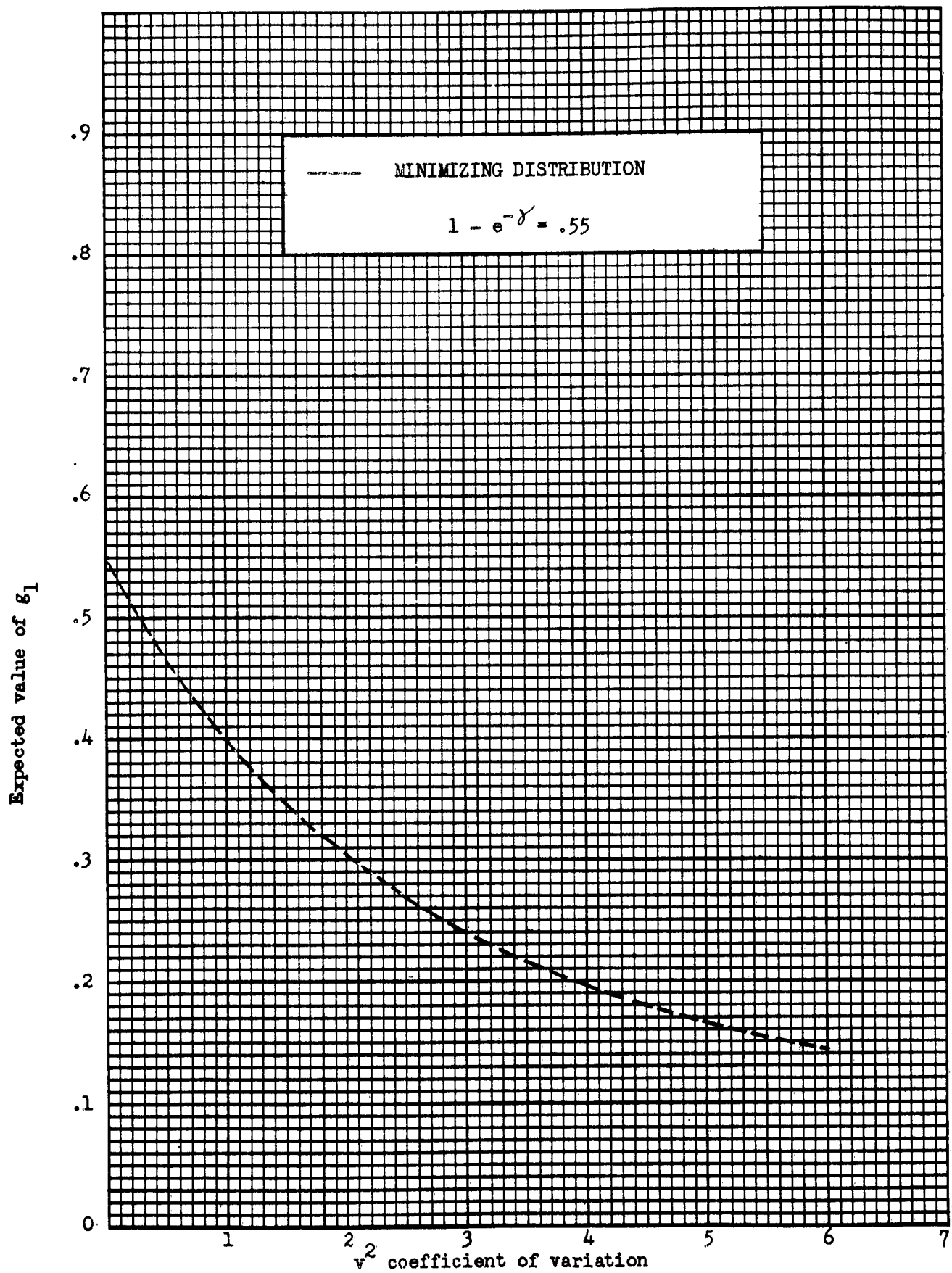


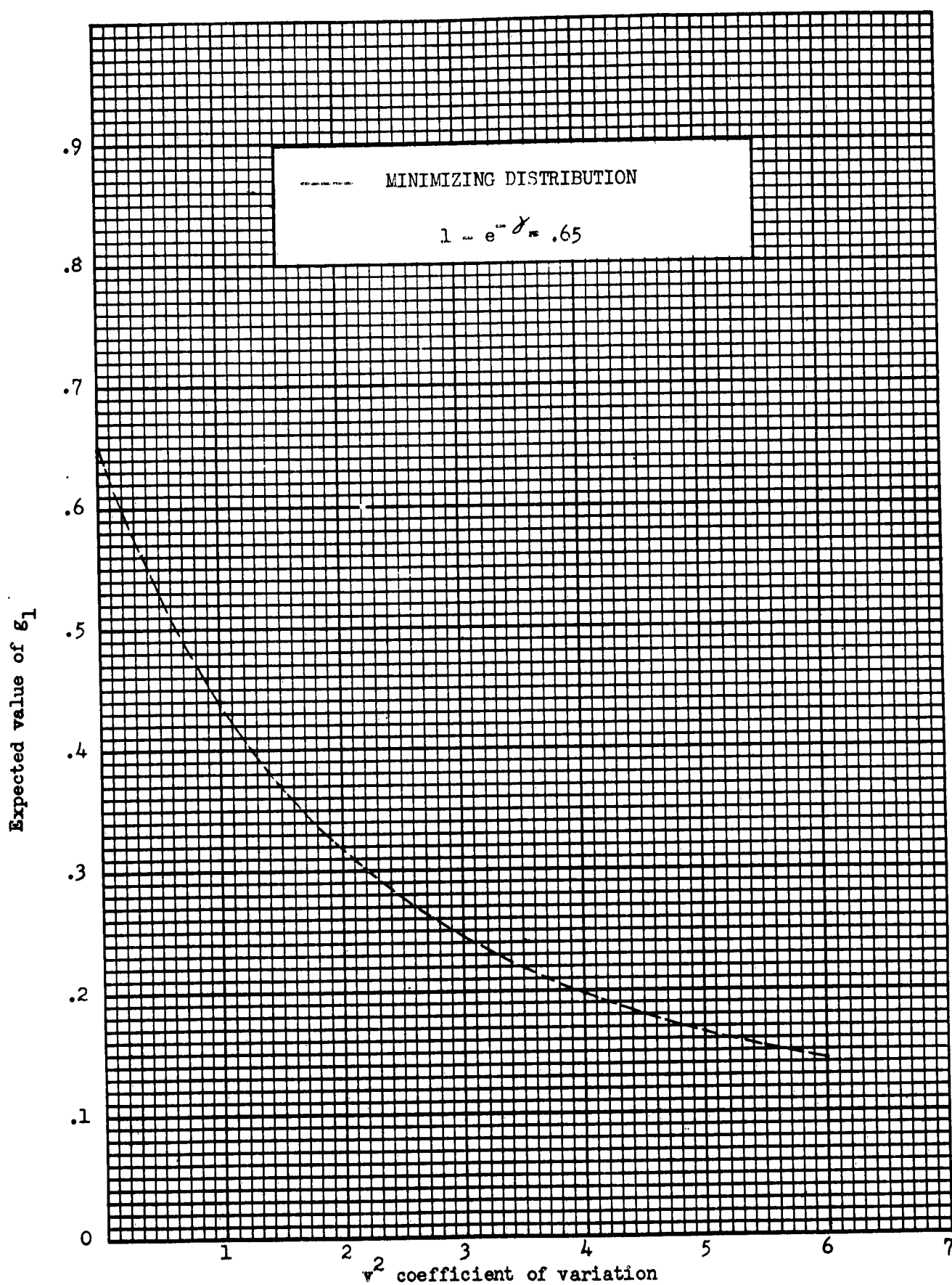


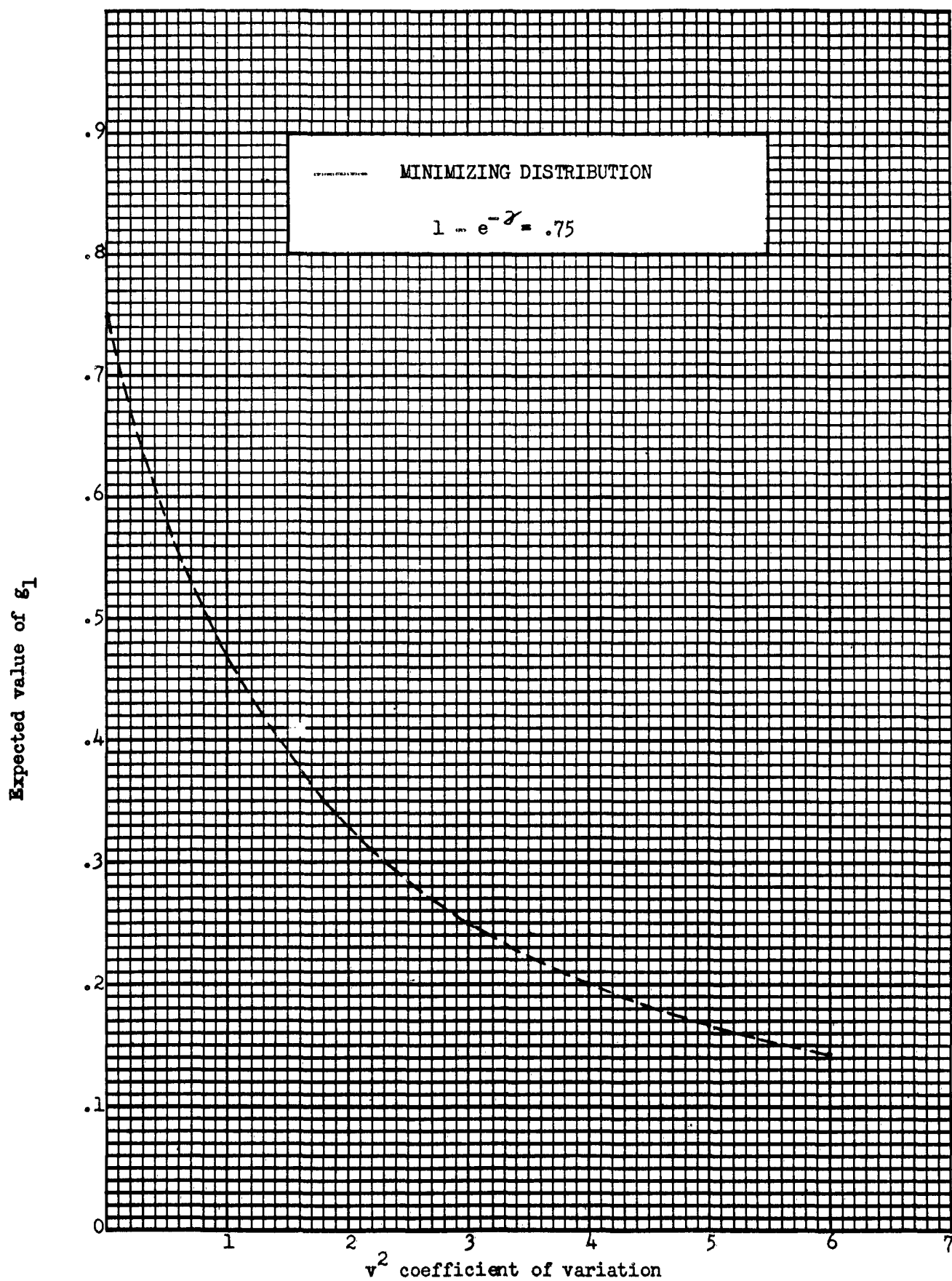


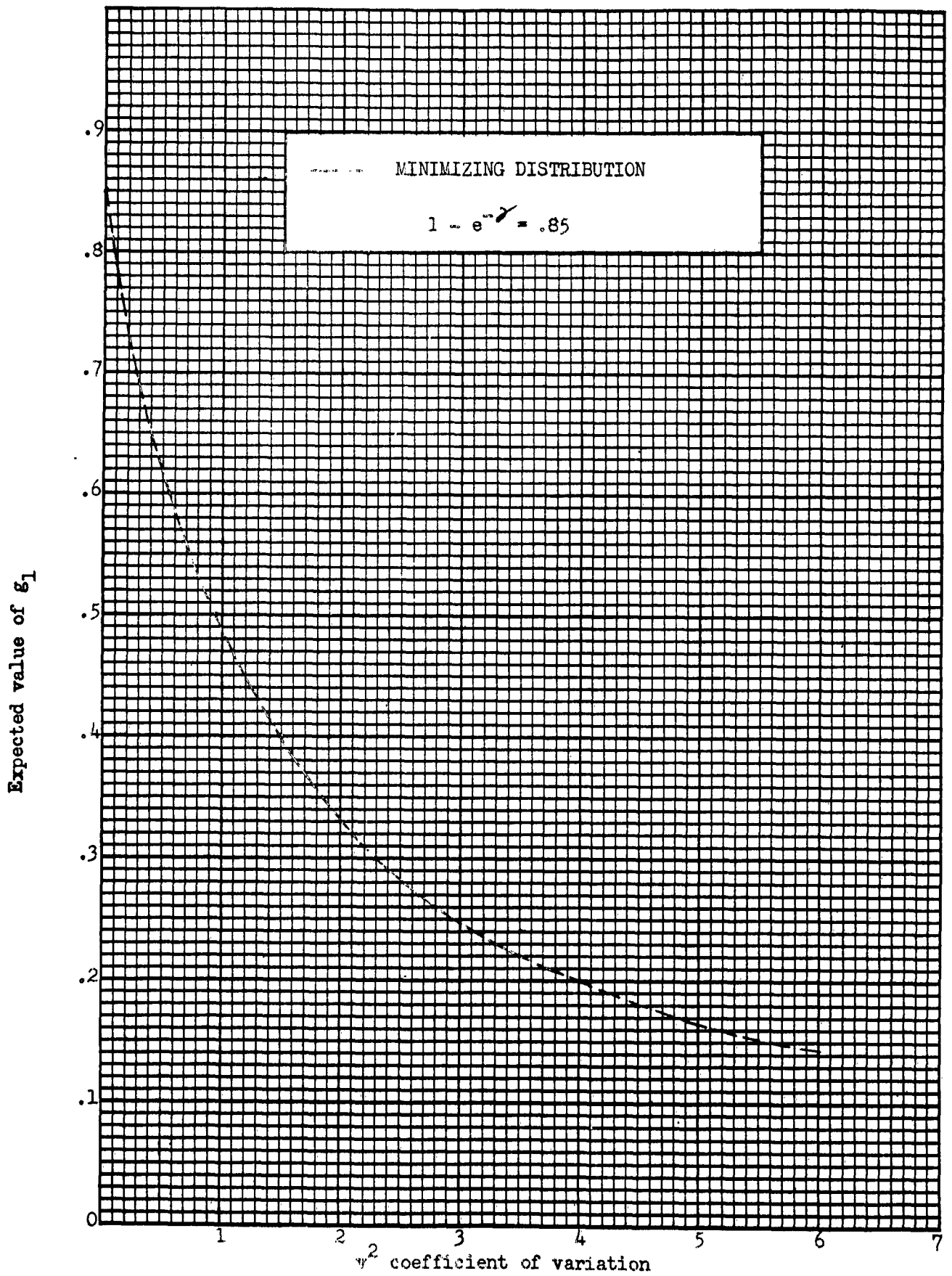


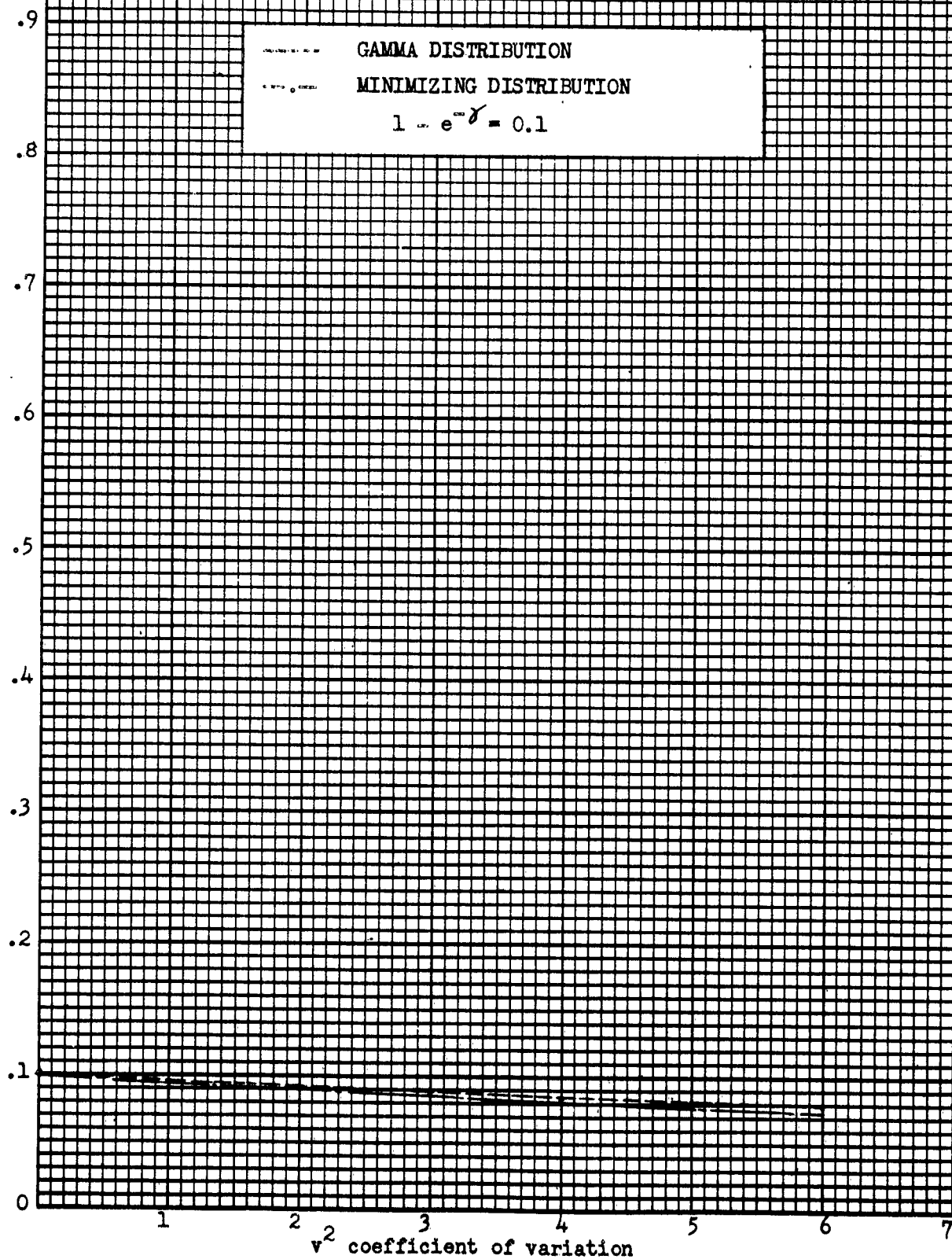


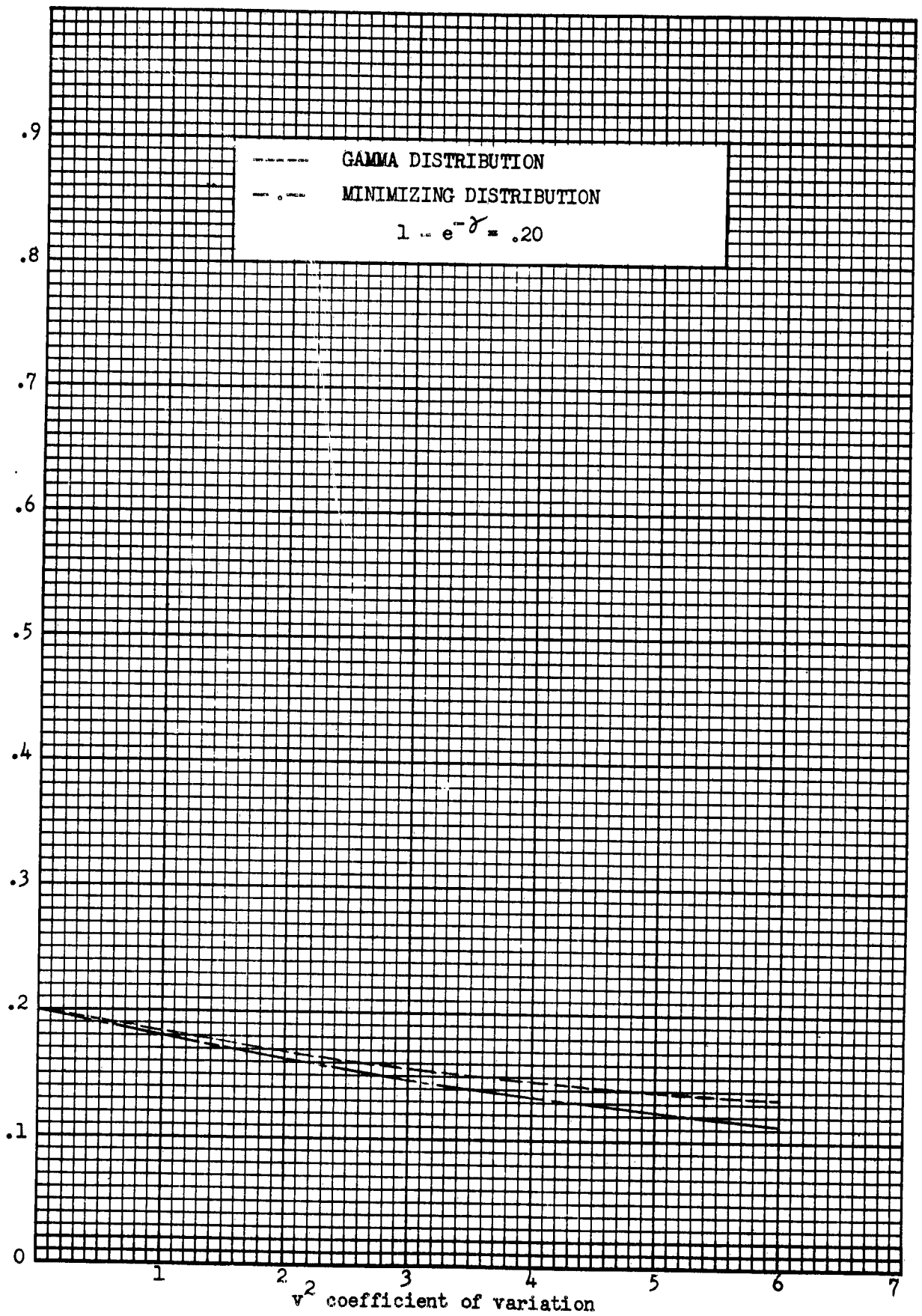


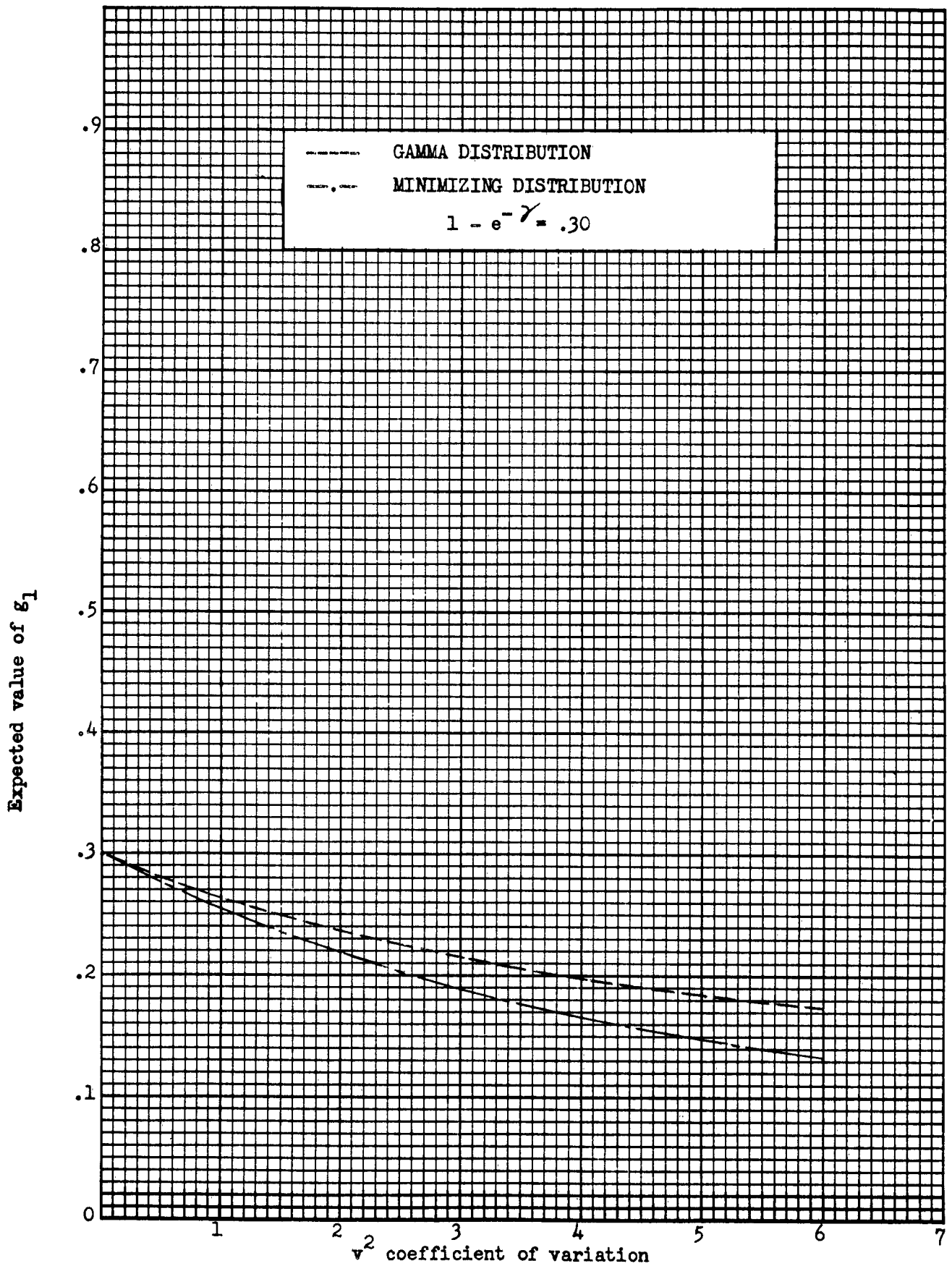


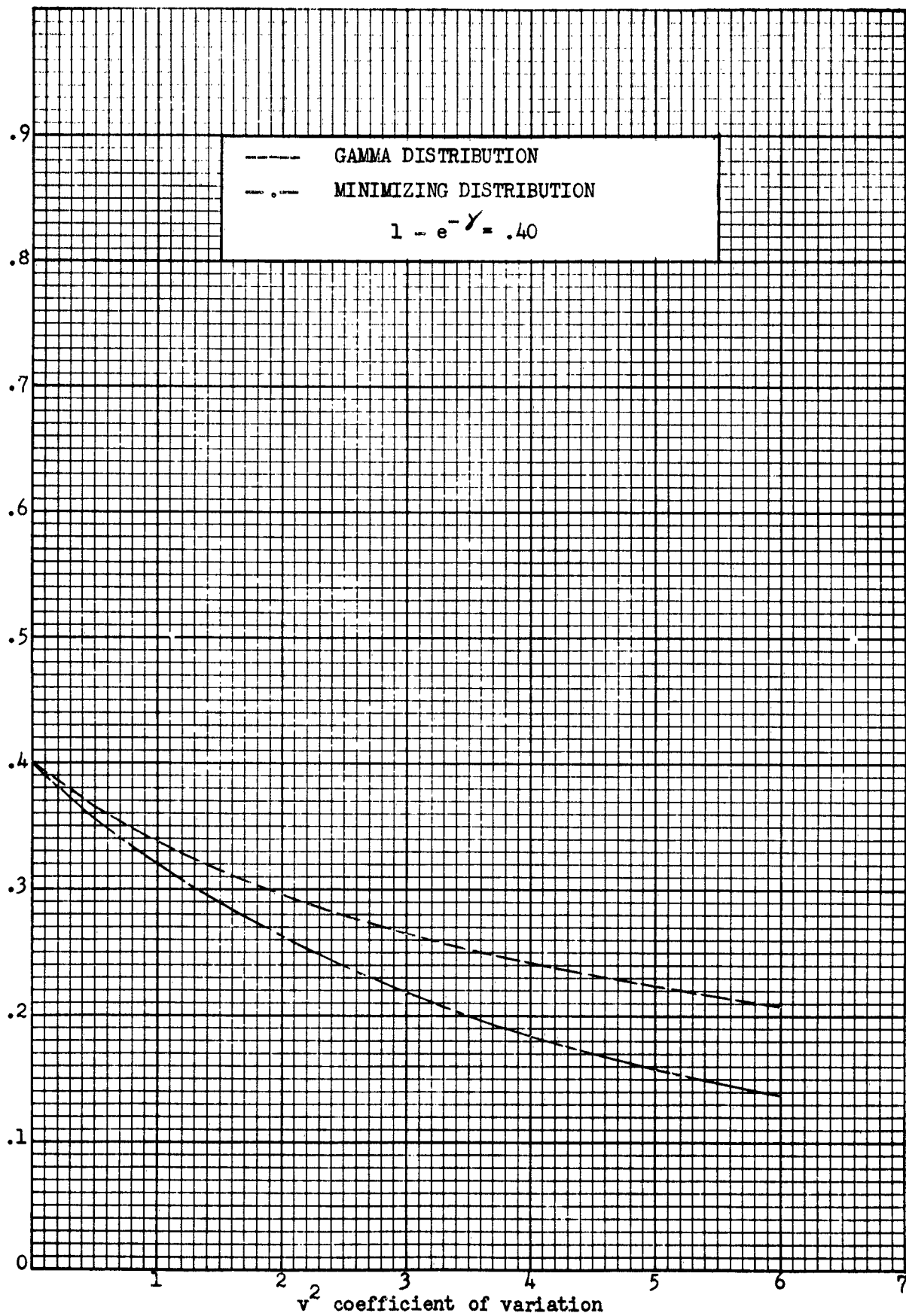


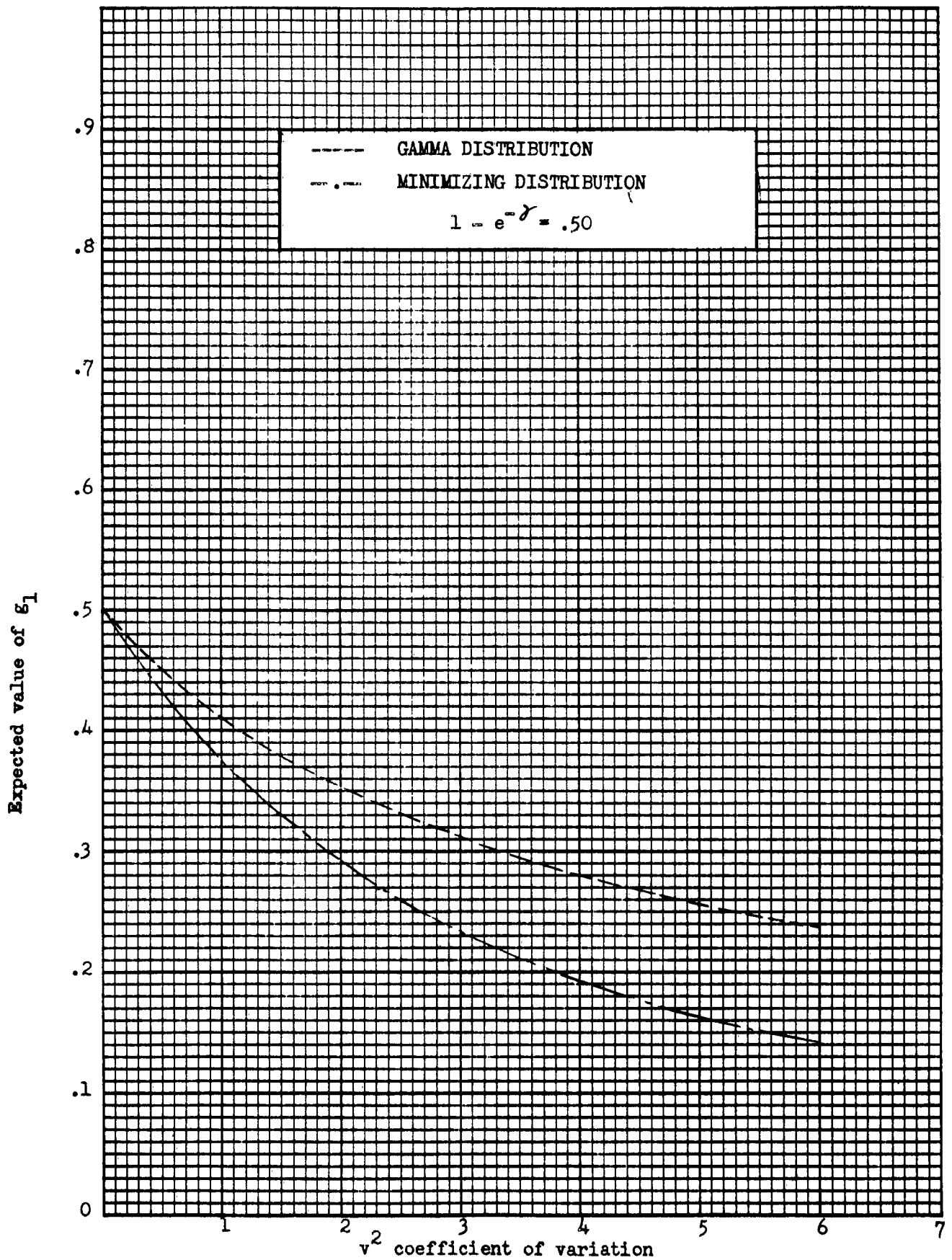


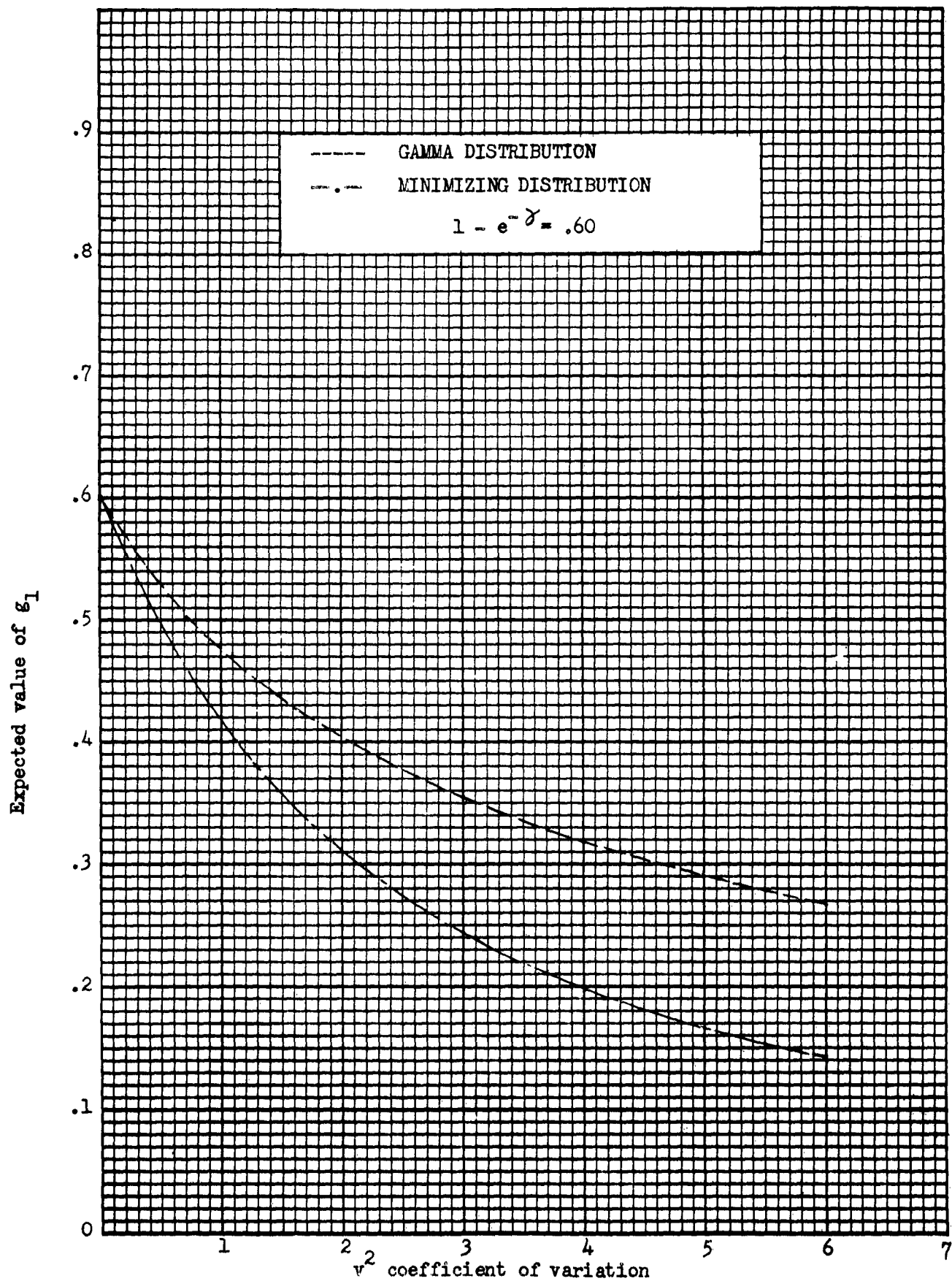
Expected value of g_1 

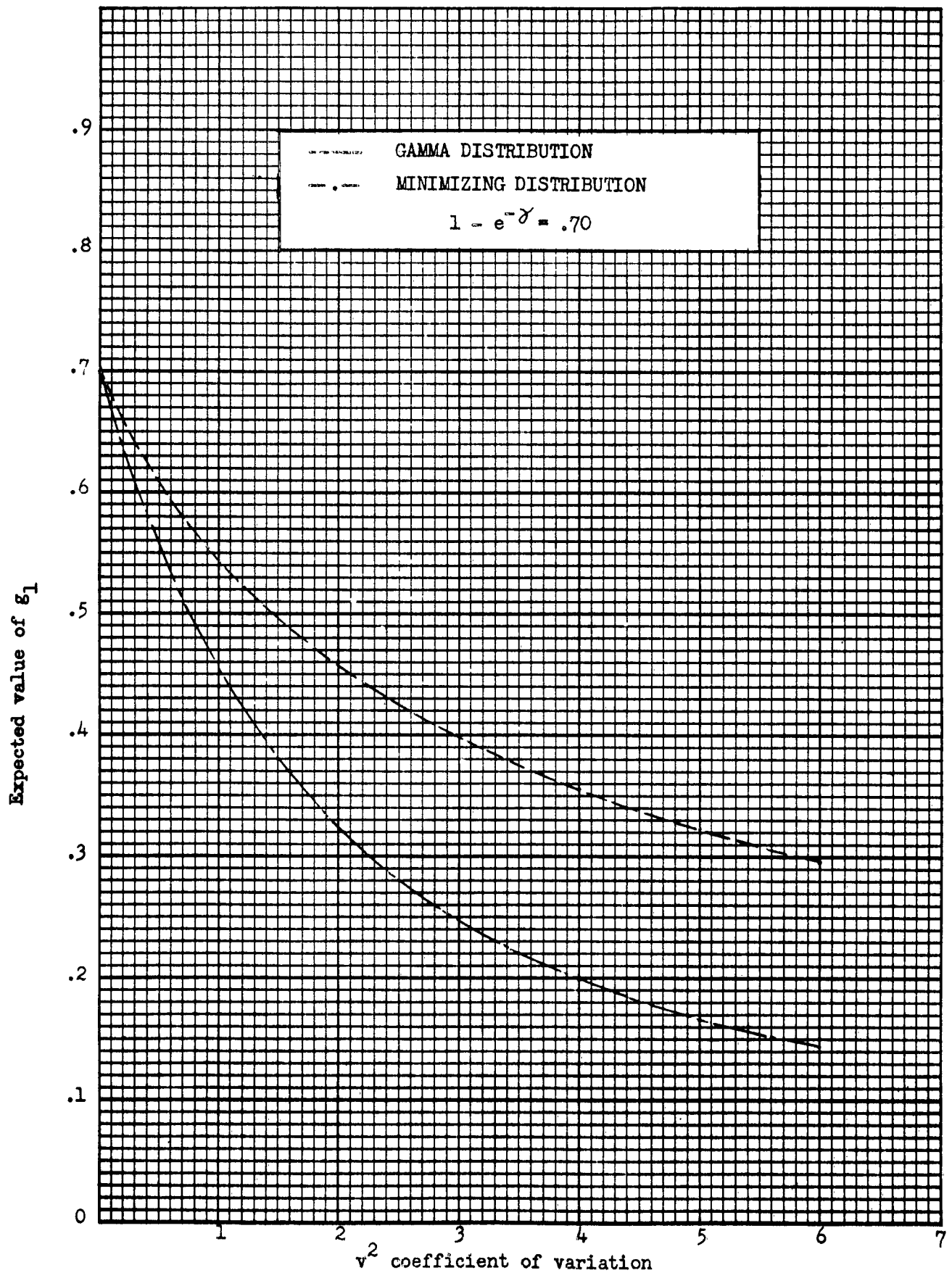
Expected value of g_1 

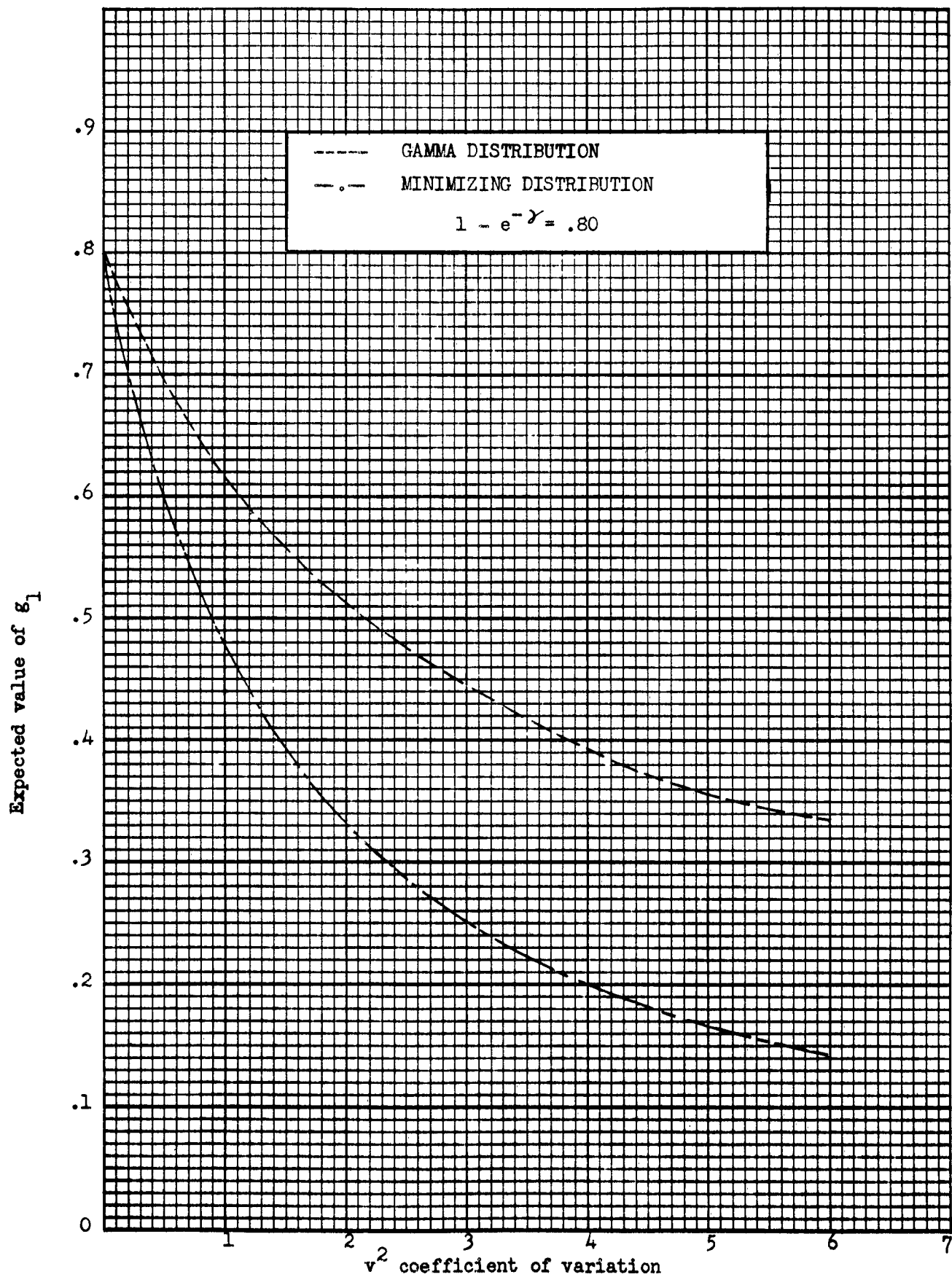


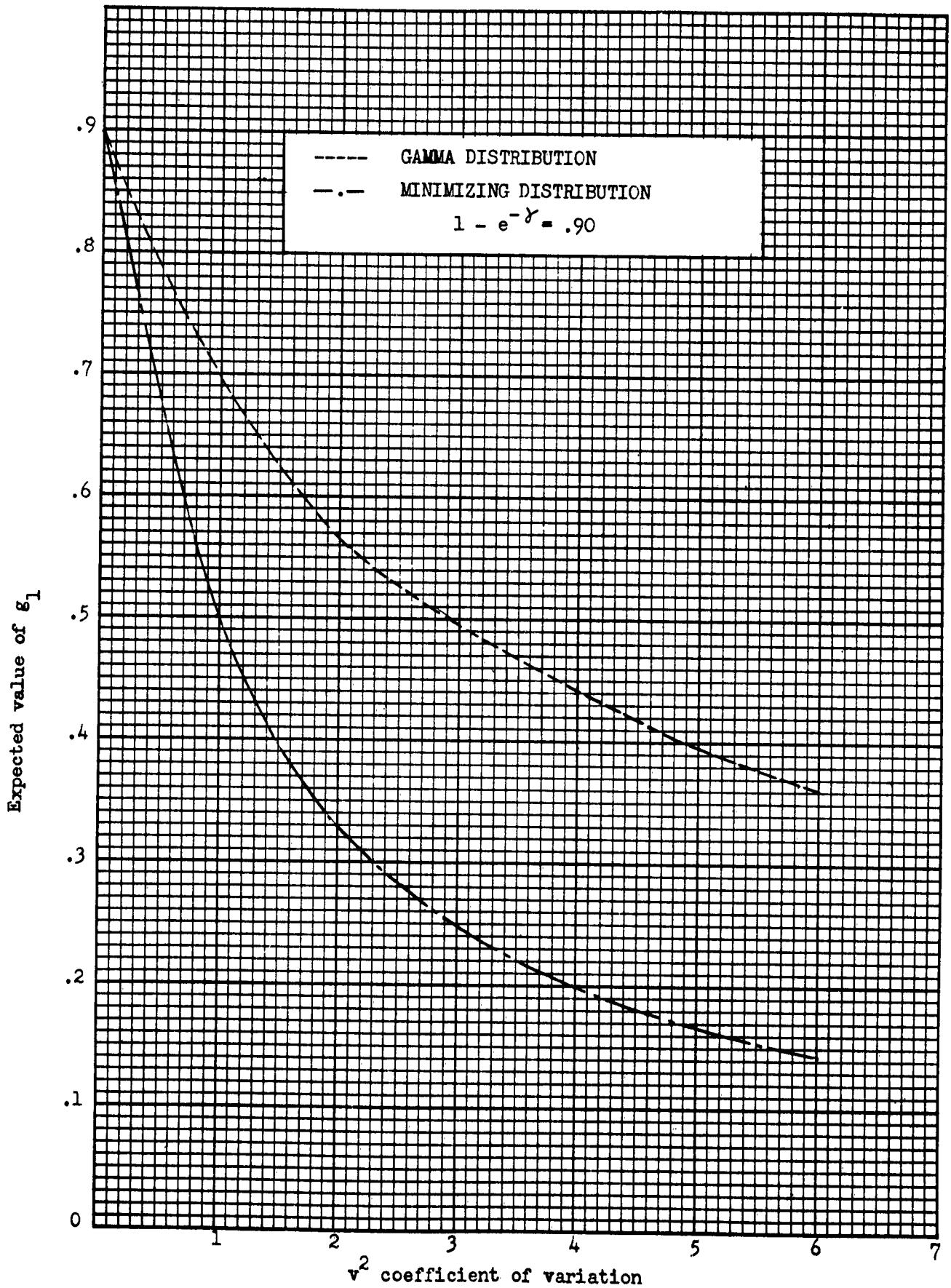
Expected value of g_1 











STANFORD UNIVERSITY

Technical Reports Distribution List

Contract N6conr 25140
(NR-342-022)

Armed Services Technical Information Agency Documents Service Center Knott Building Dayton 2, Ohio	5	Dr. Merle M. Andrew Hq. Air Research and Development Command Post Office Box 1395 Baltimore 3, Maryland Attn: RDRRM	1
Commander Wright Air Development Center Wright-Patterson Air Force Base, Ohio Attn: ARL Technical Library, WCRR	1	Prof. Z. W. Birnbaum Department of Mathematics University of Washington Seattle 5, Washington	1
Headquarters, USAF Director of Operations Washington 25, D. C. Attn: Operations Analysis Division, AFOOP	1	Dr. David Blackwell Department of Mathematics Howard University Washington 1, D. C.	1
Institute of Statistics North Carolina State College of A & E Raleigh, North Carolina	1	Prof. Albert Bowker Applied Math. and Stat. Lab. Stanford University Stanford, California	1
National Applied Mathematics Labs National Bureau of Standards Washington 25, D. C.	1	Prof. William G. Cochran Department of Biostatistics School of Hygiene and Public Health The Johns Hopkins University 615 North Wolfe Street Baltimore 5, Maryland	1
Office, Chief of Ordnance Department of the Army Washington 25, D. C. Attn: ORDTX-P	1	Dr. Walter L. Deamer, Jr. Operations Analysis Division-DCS/O Headquarters USAF Washington 25, D. C.	1
Office of Assistant Chief of Staff, G-4 Research Branch, R & D Division Department of the Army Washington 25, D. C.	1	Prof. W. J. Dixon Department of Mathematics University of Oregon Eugene, Oregon	1
Office of Technical Services Department of Commerce Washington 25, D. C.	1	Prof. M. A. Girshick Applied Math. and Stat. Lab. Stanford University Stanford, California	2
Dr. Forman S. Acton, Director Analytical Research Group Princeton University Box 710 Princeton, New Jersey	2		

Dr. I. R. Hershner Department of Mathematics University of Vermont Burlington, Vermont	1	Mr. Joseph Weinstein Applied Physics Branch Evans Signal Laboratory Fort Monmouth, New Jersey	35
Prof. Harold Hotelling Department of Mathematical Statistics University of North Carolina Chapel Hill, North Carolina	1	Prof. Jacob Wolfowitz Department of Mathematics Cornell University Ithaca, New York	1
Prof. Carl F. Kossack Department of Mathematics Purdue University Lafayette, Indiana	1		
Dr. W. W. Leutert Chief, Computing Laboratory Ballistic Research Laboratory Aberdeen Proving Ground, Maryland	1		
Dr. Eugene Lukacs Head, Statistics Branch Office of Naval Research Washington 25, D. C.	1		
Dr. Knox T. Millsaps Wright Air Development Center Wright-Patterson Air Force Base, Ohio Attn: WGRRM	1		
Prof. J. Neyman Statistical Laboratory Department of Mathematics University of California Berkeley 4, California	1		
Mr. R. H. Noyes Coles Signal Laboratory Fort Monmouth, New Jersey	1		
Prof. Herbert Robbins Department of Mathematical Statistics Columbia University New York 27, N. Y.	1		
Dr. Horace M. Trent Head, Applied Mathematics Branch Naval Research Laboratory Washington 20, D. C.	1		
Prof. W. Allen Wallis Statistical Research Center University of Chicago Chicago 37, Illinois	2		